An integrated model for the assessment of global water resources ? Part 1: Input meteorological forcing and natural hydrological cycle modules

N. Hanasaki, S. Kanae, T. Oki, K. Masuda, K. Motoya, Y. Shen, K. Tanaka

Hydrol. Earth Syst. Sci. Discuss., 4, 3535–3582, 2007
www.hydrol-earth-syst-sci-discuss.net/4/3535/2007/

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

An integrated model for the assessment of global water resources – Part 1: Input meteorological forcing and natural hydrological cycle modules

N. Hanasaki[1], S. Kanae[2], T. Oki[2], K. Masuda[3], K. Motoya[4], Y. Shen[5], and K. Tanaka[6]

[1]National Institute for Environmental Studies, Japan
[2]Institute of Industrial Science, The University of Tokyo, Japan
[3]Frontier Research Center for Global Change, Japan
[4]Faculty of Education and Human Studies, Akita University, Japan
[5]Center for Agricultural Resources Research, The Chinese Academy of Sciences, China
[6]Disaster Prevention Research Institute, Kyoto University, Japan

Received: 17 September 2007 – Accepted: 17 September 2007 – Published: 2 October 2007

Correspondence to: N. Hanasaki (hanasaki@nies.go.jp)

Abstract

An integrated global water resources model was developed consisting of six modules: land surface hydrology, river routing, crop growth, reservoir operation, environmental flow requirement estimation, and anthropogenic water withdrawal. It simulates both natural and anthropogenic water flow globally (excluding Antarctica) on a daily basis at a spatial resolution of 1°×1° (longitude and latitude). The simulation period is 10 years, from 1986 to 1995. This first part of the two-feature report describes the input meteorological forcing and natural hydrological cycle modules of the integrated model, namely the land surface hydrology module and the river routing module. The input meteorological forcing was provided by the second Global Soil Wetness Project (GSWP2), an international land surface modeling project. Several reported shortcomings of the forcing component were improved. The land surface hydrology module was developed based on a bucket type model that simulates energy and water balance on land surfaces. Simulated runoff was compared and validated with observation-based global runoff data sets and observed streamflow records at 32 major river gauging stations around the world. Mean annual runoff agreed well with earlier studies at global, continental, and continental zonal mean scales, indicating the validity of the input meteorological data and land surface hydrology module. In individual basins, the mean bias was less than ±20% in 14 of the 32 river basins and less than ±50% in 24 of the basins. The performance was similar to the best available precedent studies with closure of energy and water. The timing of the peak in streamflow and the shape of monthly hydrographs were well simulated in most of the river basins when large lakes or reservoirs did not affect them. The results indicate that the input meteorological forcing component and the land surface hydrology module provide a framework with which to assess global water resources, with the potential application to investigate the subannual variability in water resources. GSWP2 participants are encouraged to re-run their model using this newly developed meteorological forcing input, which is in identical format to the original GSWP2 forcing input.

HESSD

4, 3535–3582, 2007

1 Introduction

Water is one of the most fundamental resources for humans and society. Rapid growth of the world population and economy has brought major increases in water demand during the 20th century, and this trend is projected to continue in the 21st century (Shiklomanov, 2000). Several global water resource assessments were released to project the current and future distributions of water-deficient areas worldwide (Arnell, 1999, 2004; Vörösmarty et al., 2000; Oki et al., 2001, 2003; Alcamo et al., 2003a, b, 2007; Oki and Kanae, 2006). These assessments used global hydrological models to estimate the distribution of runoff (renewable freshwater) and various world statistics to estimate water use. A typical approach is to display the global distribution of per capita annual water resources or the ratio of withdrawal to water resources on an annual basis. However, extreme seasonality in both water resources and water use occurs in some parts of the world. For example, in the Asian monsoon region, conditions change dramatically between the rainy and dry seasons. Moreover, global warming is projected to alter future temperature and precipitation patterns and consequently affect both amount and timing of water resource availability and use. Therefore, subannual variability must be taken into account.

A model suitable for such assessments requires the following three functions. First, it must simulate both renewable freshwater resources and water use at a subannual timescale. Second, it must deal with major interactions between the natural hydrological cycle and anthropogenic activities. For example, withdrawal from the upper stream affects availability in the lower stream, and reservoir operation may contribute to increased water availability in the lower stream. Third, it must explain key mechanisms regarding the effects of global warming on water resources and water use for future projections. In this two-feature report, we introduce an integrated global water resources model and an assessment of global water resources through the application of the model.

The integrated global water resources model consists of six modules: land surface

hydrology, river routing, crop growth, reservoir operation, environmental flow requirements, and anthropogenic water withdrawal. It simulates both natural and anthropogenic water flow globally (excluding Antarctica) at a spatial resolution of 1°×1° (longitude and latitude) at a daily time interval.

Several integrated global water resources models that can simulate not only the natural water cycle, but also anthropogenic water flow, have been published. Alcamo et al. (2003a, b) developed a global water assessment model called "WaterGAP 2," which consists of a global water use model and a global hydrology model, and assessed the current and future water resources globally. The global water use model consists of domestic and industry sectors, which account for the effects of structural and technological changes on water use, and the agriculture sector, which accounts especially for the effect of climate on irrigation water requirements. The global hydrology model calculates surface runoff and groundwater recharge based on the computation of the daily water balance of the soil and canopy. Water balance is also estimated for surface waters, and river flow is regulated via a global flow routing scheme. Major human-made reservoirs of the world were geo-referenced to the river network, but their outflow was identical to that of natural lakes. Haddeland et al. (2006) developed and implemented a reservoir model and an irrigation model in the Variable Infiltration Capacity (VIC) land surface model (Liang et al., 1994) and studied the effects of reservoirs and irrigation water withdrawal on continental surface water fluxes for part of North America and for Asia. Jachner et al. (2007) enhanced the LPJmL (Lund-Potsdam-Jena managed land) dynamic global vegetation model (Bondeau et al., 2007) with a river routing model, including lakes and reservoirs, and withdrawals for households and industry and assessed how much water is consumed by global irrigated and rain-fed agriculture and by natural ecosystems.

We set two basic policies. First, both water and energy balances on the land surface are closed in each module and the integrated model. This is not only the most fundamental consideration of hydrology, but is also one of the key requirements for the interdisciplinary coupling of submodules (e.g., hydrological models and crop growth

HESSD
4, 3535–3582, 2007

models). Recently, several advanced earth system modeling efforts have been reported such as coupling a land surface model (LSM) with a crop model (Gervois et al., 2004; Mo et al., 2005), and coupling a global climate model (GCM) with a crop model (Osborne et al., 2007). In these systems, soil moisture, evaporation, and other variables are shared by more than one submodel; therefore, to maintain consistency among submodels, energy and water balances should be conserved. In particular, the closure of the energy balance is a fundamental requirement of GCM and LSM approaches. Second, we tried to avoid model calibration involving the fit of simulated results to available observation records. It is well established that hydrological models do not reproduce observed hydrographs very well without model calibration (or model parameter tuning). However, in global-scale hydrological modeling, model calibration is a difficult issue. There are a few reasons for this. First, it is virtually impossible to calibrate the model worldwide because of the limited availability of observations, especially in developing countries. Second, both models and input meteorological forcing and validation data contain considerable uncertainty (Oki et al., 1999), and it is not always easy to attribute errors in simulations to improper settings of model parameters. Moreover, we intended to apply the model to future projection under climate change. Thus, the transparency and physical validity of the model are quite important because the simulated results are highly model dependent. Therefore, we extensively examined the simulated results of the model using model inherent parameters; even this sometimes produces large errors.

The simulation was conducted using the framework of the second Global Soil Wetness Project (GSWP2; Dirmeyer et al., 2006), which is an international project that estimated the global energy and water balance over land, with emphasis on variation in soil moisture. This framework has two significant benefits. First, it provides quality-checked input meteorological forcing (e.g., air temperature and precipitation) and consistent surface boundary conditions (e.g., land-sea mask and albedo) with which to simulate energy and water balances globally. Second, it allows the comparison of our model with state-of-the-art land surface models involved in the GSWP2.

HESSD
4, 3535–3582, 2007

Here, we describe the input meteorological forcing and natural hydrological cycle modules (i.e., land surface hydrology and river routing). In modeling and simulations, the preparation of reliable meteorological forcing inputs is essential. First, we revisited the GSWP2 activities and the original meteorological forcing inputs of the GSWP2 (Sect. 2). We traced some of its shortcomings and developed improved meteorological forcings (Sect. 3). We then described two natural hydrological modules and simulation settings (Sects. 4, 5, 6). We compared our simulated results with observations and earlier studies, and the performance of the input meteorological data and modules was validated (Sect. 7). The results demonstrated that the natural hydrological modules can reproduce the global, continental, and continental zonal mean distributions of runoff. Finally, the mean annual streamflow, timing of peak streamflow, and interannual fluctuation in annual streamflow in 32 major river basins from around the world were validated. The results indicated that their reproducibility rivals the best available earlier studies. In a forthcoming paper (Hanasaki et al., 2007), we describe and validate the anthropogenic activities modules of crop growth, reservoir operation, environmental flow requirements, and anthropogenic water withdrawal. All six modules are then coupled in an integrated model, and global water resource assessments are conducted.

Here, "runoff" indicates the water that drains from surfaces and subsurfaces of a certain area of land [mm yr^{-1} or mm mo^{-1}], and "streamflow" indicates the flow of water in river channels [m^3 s^{-1}].

2 Global Soil Wetness Project

The Global Soil Wetness Project (GSWP) plays an important role in understanding and quantifying the global hydrological cycle (Dirmeyer et al., 1999, 2006). The GSWP is an international scientific project that aims to produce state of the art global data on land surface fluxes (e.g., sensible heat, latent heat, and runoff) and state variables (e.g., soil moisture and soil temperature). It is a modeling project that is used to drive dozens of state-of-the-art land surface models (LSMs) in offline mode with common meteoro-

logical forcing and boundary conditions from the International Satellite Land Surface Climatology Project (ISLSCP; Sellers et al., 1996; Hall et al., 2006). The GSWP also aims for the intercomparison of participating LSMs and large-scale validation and quality checks of ISLSCP because there are uncertainties in both meteorological forcing inputs and land surface models.

In the first phase of the GSWP (hereafter GSWP1; Dirmeyer et al., 1999), 11 state-of-the-art LSMs were driven in offline mode using global meteorological forcing inputs, vegetation cover, and soil cover provided by the ISLSCP initiative 1 (ISLSCP1; Meeson et al., 1995, Sellers et al., 1996). The GSWP1 succeeded in producing comprehensive global land flux and state variables data for 1987–1988 at a spatial resolution of 1°×1° at 10-day intervals. Subsequently, intensive validation of output products and model intercomparisons were conducted. For runoff, the major shortcomings of the GSWP1 were its short simulation period and its tendency for underestimation. A 2-year period is too short for the discussion of interannual variation in runoff and long-term mean runoff. The annual mean total global runoff of GSWP1 (the model ensemble mean of 11 participating LSMs) is 29 485 $km^3\,yr^{-1}$ (Oki et al., 2001), whereas a large number of earlier studies reported it to be approximately 40 000 $km^3\,yr^{-1}$ (Baumgartner and Reichel, 1975; Fekete et al., 2002, and Döll et al., 2003). Oki et al. (1999) pointed out that significant underestimation in runoff is observed at higher latitudes and is attributable to the quality of precipitation forcing inputs and rain gauge undercatch in strong winds.

Recently, the second phase of the Global Soil Wetness Project (GSWP2; Dirmeyer et al., 2006) was conducted. In the GSWP2, 15 LSMs participated and produced daily land flux and state variable data for 10 years (1986–1995) at a spatial resolution of 1°×1°. The meteorological forcing input, vegetation cover, and soil cover data were provided primarily by the ISLSCP Initiative 2 (ISLSCP2; Hall et al., 2006), and elaborate work to make the data usable as meteorological forcing inputs was undertaken by the Center for Ocean-Land-Atmosphere Studies (COLA; Zhao and Dirmeyer, 2003). Taking the findings of Oki et al. (1999) into account, a rain gauge undercatch correction was applied to the precipitation forcing data. Therefore, the major shortcomings of the

HESSD
4, 3535–3582, 2007

GSWP1 for runoff, i.e., short simulation period and underestimation in runoff, were expected to be reduced in the GSWP2.

Some preceding studies have validated the output of the GSWP2 using field observations and evaluated the meteorological forcing inputs and the performance of the participating LSMs. Guo and Dirmeyer (2006) and Guo et al. (2006) validated the GSWP2 soil moisture results against field observations. They found that participating land surface models reproduced the observed interannual variability and phasing of the soil moisture annual cycle reasonably well. Decharme and Douville (2006) validated the precipitation forcing input and runoff output of the GSWP2 at the Rhône River basin, where reliable observations are available, and at 80 gauging stations distributed over the globe. They showed that the runoff of the GSWP2 was overestimated at middle and high latitudes and attributed it to the precipitation forcing input. These findings indicate that the meteorological forcing input of the GSWP2 needs to be revisited.

3 Meteorological forcing input

In this section, we revisit the original meteorological forcing input of the GSWP2 (hereafter F-GSWP2-B0; F stands for forcing, B0 for baseline experiment version zero). In the GSWP2, seven input meteorological components were provided to drive the LSMs, namely, downward longwave radiation, downward shortwave radiation, wind speed, surface air pressure, specific humidity, air temperature, and precipitation. All of these components were provided for 10 years (1986–1995) at 3-h intervals, covering all land excluding Antarctica at a spatial resolution of $1^{\circ} \times 1^{\circ}$. The methodology for producing these components is described in detail by Zhao and Dirmeyer (2003), and a short description can be found in Appendix A. F-GSWP2-B0 is a hybrid product of the NCEP-DOE reanalysis (Kanamitsu et al., 2002) and various observation-based monthly meteorological data on a global grid (see Table 1). The NCEP-DOE reanalysis was corrected linearly to match the monthly mean values to the observation-based data. For the precipitation data, an algorithm for the gauge correction of wind-caused

undercatch was applied to the rainfall and snowfall input data (Motoya et al., 2002). In Motoya's algorithm, the catchment ratio of snowfall (CR_{snow}) and that of rainfall (CR_{rain}) are expressed as follows:

$$CR_{\text{snow}} = \begin{cases} a\exp(bU) & \text{raingauge type known} \\ 50.0\exp(-0.182U) + 50.0\exp(-0.112U) & \text{raingauge type unknown} \end{cases} \quad (1)$$

$$CR_{\text{rain}} = 100.0 - 1.51U - 0.21U^2 \quad (2)$$

where U is wind speed at a height of 2 m, and a and b are parameters that depend on the rain gauge type from Sevruk (1982) and Sevruk and Hamon (1984). The corrected rainfall ("Rainf") and snowfall ("Snowf") are expressed as:

$$\begin{cases} \text{Snowf} = \text{Snowf}_{\text{org}} / CR_{\text{snow}} \\ \text{Rainf} = \text{Rainf}_{\text{org}} / CR_{\text{rain}} \end{cases} \quad (3)$$

where $\text{Rainf}_{\text{org}}$ and $\text{Snowf}_{\text{org}}$ are the original rainfall and snowfall, respectively. These equations indicate that stronger wind brings a stronger correction and snowfall requires a stronger correction than rainfall. For example, if the wind speed is 5 m s^{-1} and the rain gauge type is unknown, the CR_{snow} is 48.7%, and CR_{rain} is 87.2%.

To show the characteristics of F-GSWP2-B0, the global zonal mean distributions of wind speed and precipitation are provided (Fig. 1). As a yardstick, the mean 1961–1990 global observation-based data of the Climate Research Unit at the University of East Anglia (CRU; New et al., 1999) are shown for wind speed and precipitation. The precipitation of F-GSWP2-B0 is clearly larger than that of the CRU at middle to high latitudes, but smaller at low latitudes. Decharme and Douville (2006) argued that high precipitation at middle to high latitudes caused the overestimation of runoff, which is commonly observed in all GSWP2 participating LSMs. The wind speed of F-GSWP2-B0 is much higher than that of the CRU, except at low latitudes. This is one possible cause of overestimation in precipitation of F-GSWP2-B0 because Motoya's undercatch correction is correlated with wind speed (Eqs. 1 and 2), especially in regions at high latitudes in which precipitation is dominated by snow.

We obtained the original source program code that was used to produce F-GSWP2-B0 precipitation data and found that wind speed at a height of 10 m was used, whereas Motoya's algorithm expects a height of 2 m. This is another cause of overcorrection in F-GSWP2-B0 precipitation data because wind speed is stronger at higher altitudes.

To correct and improve the accuracy of F-GSWP2-B0, a new input meteorological data set (F-GSWP2-B1) was developed. The major differences included a change in reanalysis data from NCEP-DOE (Kanamitsu et al., 2002) to ERA40 (Betts and Beljaars, 2003) and the proper application of Motoya's wind correction to precipitation data. The differences in each variable between the two forcing data sets were determined (Table 1). After the completion of F-GSWP2-B0, the European Centre for Medium Range Weather Forecast (ECMWF) near-surface meteorological data set for the ISLSCP2 data collection was released (ERA40, Betts and Beljaars, 2003). Tanaka et al. (2005) compared air temperature, vapor pressure, wind speed, and precipitation of NCEP-DOE and ERA40 with daily ground meteorological observations collected by the National Climatic Data Center (http://www.ncdc.noaa.gov) for 1994–1995 at 2349 stations around the world and found better agreement with ERA40 than with NCEP-DOE. Because daily and diurnal variation in the GSWP2 meteorological forcing inputs is dependent on the reanalysis data, we substituted the ERA40 data for the NCEP-DOE data.

The wind speed of F-GSWP2-B1 is more similar to that of the CRU than the F-GSWP2-B0, but it is smaller than that of the CRU at southern low latitudes (Fig. 1). For precipitation, F-GSWP2-B1 has greater precipitation at latitudes higher than 50° N in the Northern Hemisphere and higher than 35° S in the Southern Hemisphere, but the difference from the CRU is much smaller than that from the F-GSWP2-B0 (Fig. 1).

4 Model

4.1 Land surface hydrology module

A land surface hydrology module calculates the energy and water budget, including snow, on the land surface from the forcing data. This module is based on a bucket model (Manabe, 1969; Robock et al., 1995), but differs from the original formulation in the following three aspects. First, soil temperature is calculated using the force restore method (Bhumralkar, 1975; Deardorff, 1978) to simulate the diurnal cycle of surface temperature reasonably using three-hourly meteorological forcing inputs. Second, a simple subsurface runoff parameterization is added to the model. Third, two independent land surface conditions can be simulated within a single grid that is intended to separate irrigated cropland from other land types. This function is not used here, but in the companion paper. The bucket model is simple, but is still widely used in current global water hydrological studies. Soil moisture is expressed as a single-layer bucket 15 cm deep for all soil and vegetation types. When the bucket is empty, soil moisture is at the wilting point; when the bucket is full, soil moisture is at field capacity. Evapotranspiration is expressed as a function of potential evapotranspiration and soil moisture. In the original bucket model, runoff is generated only when the bucket is overfilled, but we used a "leaky bucket" formulation in which soil moisture drains continuously. Potential evapotranspiration and snowmelt are calculated from the surface energy balance. A detailed description of this module can be found in Appendix B.

4.2 River routing module

The river module is identical to the Total Runoff Integrating Pathways model (TRIP; Oki and Sud, 1998; Oki et al., 1999). The module has a digital river map with a spatial resolution of $1^{\circ} \times 1^{\circ}$ (the land–sea mask is identical to the GSWP2 meteorological forcing input) and a flow velocity fixed at $0.5\,\mathrm{m\,s^{-1}}$. The module accumulates runoff calculated by the land surface hydrology module and outputs streamflow. This mod-

HESSD
4, 3535–3582, 2007

ule does not deal with lakes or swamps, human-made reservoir operation, diversion or withdrawal, or evaporation loss from water surfaces. Human-made reservoir operation and withdrawal are modeled in the anthropogenic activity modules.

5 Observations and earlier studies

Observed streamflow data were obtained from Global Runoff Data Centre (GRDC). From the 3045 gauging stations available, the 37 stations used here have catchment area >200 000 km^2, monthly streamflow records for more than 60 months between January 1986 and December 1995, and the largest number of downstream river gauging stations in their respective basins. Of the 37 stations, five river basins in arid zones ($R_{\mathrm{net}}/\iota P$>0.2, see Sect. 6.2 and Fig. 3; Niger River, Darling River, Orange River, Colorado River, and Cooper Creek) were excluded because most previous studies significantly overestimate observation records from these basins. Thus, we used 32 stations and their catchment areas (Fig. 2).

To compare simulated runoff with that in previous studies, four published runoff data sets, namely Baumgartner and Reichel (1975), Nijssen et al. (2001), Fekete et al. (2002), and Döll et al. (2003) were collected (Table 2). The global runoff data set of Baumgartner and Reichel (1975) is often cited as a benchmark global hydrological cycle study (hereafter R-BR75). They overlaid a number of maps of precipitation, evaporation, and runoff and corrected them until water was balanced. They provided mean 5° zonal mean precipitation, evaporation, and runoff data for each continent.

Nijssen et al. (2001) developed a 14-year (1980–1993) data set of global gridded runoff (hereafter R-N01) by driving the Variable Infiltration Capacity (VIC) land surface model (Liang et al., 1994) in offline mode. In contrast to the GSWP, they tuned the model parameters for 26 rivers in five climatic zones (eight climatic zones after subdivision). For the remaining river basins, parameter sets from the nearest tuned river basin in the same climatic zone were applied. They showed that this parameter transfer decreased the mean bias of annual streamflow and the root mean square error (RMSE)

HESSD
4, 3535–3582, 2007

of monthly streamflow.

Fekete et al. (2002) developed a global gridded runoff data set (hereafter, R-F02) by interpolating 663 major river gauging stations from the GRDC. First, they georeferenced each river gauging station to a global digital river map (0.5°×0.5°) and calculated the mean annual interstation runoff. A macro-scale hydrological model was then driven and global gridded monthly runoff was simulated. Their model required only two meteorological forcing inputs (monthly air temperature and precipitation), and energy balance was not solved. Finally, the simulated runoff was corrected at every interstation basin using an adjustment multiplier defined as the ratio of observed interstation runoff to simulated runoff. For basins lacking observations, the simulated runoff was used without any correction. These runoff data are climatological, and the period is not specified.

Döll et al. (2003) developed the Water GAP Hydrological Model and produced global gridded monthly runoff data for 95 years (1901–1995; hereafter R-D03). Their model requests monthly climate inputs provided by New et al. (2000), namely air temperature, precipitation, number of wet days per month, cloudiness, and average daily sunshine hours. These variables were interpolated temporally to create daily data. The model had one tuning parameter that was tuned at 724 GRDC river gauging stations to reproduce the observed long-term mean annual streamflow. The parameter varied within a certain physically plausible range. When the parameter exceeded the range (339 of 724 basins), a correction similar to that of Fekete et al. (2002) was applied.

Of these four data sets, R-BR75, R-D03, and R-F02 are regarded as observation-based runoff products. They report global annual runoff of approximately 280 mm yr^{-1} (36 500 km^3 yr^{-1}, excluding Antarctica). R-D03 accounts for anthropogenic water withdrawal, and 9.5 mm yr^{-1} (1250 km^3 yr^{-1}) of consumptive water use is extracted from its natural runoff.

6 Simulations

6.1 Global natural hydrological simulation

The land surface hydrology module was driven using three sets of GSWP meteorological forcing input, namely F-GSWP1, F-GSWP2-B0, and F-GSWP2-B1, and three global gridded runoff products were obtained: R-GSWP1, R-GSWP2-B0, and R-GSWP2-B1, respectively. The simulation period of R-GSWP1 was 2 years (1987–1988) and that of R-GSWP2-B0 and R-GSWP2-B1 was 10 years (1986–1995). The land–sea mask and albedo were not identical in each project. The runoff product was routed using the river routing module.

6.2 Parameter modification

We compared R-GSWP2-B1 with earlier studies and validated it using observed streamflow; there were substantial regional biases. The parameter settings of the land surface hydrology module, which were set uniformly over the globe, appeared to be responsible (see Appendix B). An analysis of energy and water constraint provided insights into the characteristics of the land surface hydrology module. For 37 basins (including five semi-arid and arid river basins that were excluded from the validation), Budyko's aridity index (the ratio of net radiation to precipitation, $R_{\mathrm{net}}/\iota P$; Budyko, 1974) and the evaporation to precipitation ratio ($E/P=(P-R)/P$) were calculated and plotted in a Budyko diagram (Fig. 3a). To investigate the relationship between Budyko's diagram and the climatic zone, all land grid cells were initially classified using Köppen's climate classification (McKnight and Hess, 2000), using the monthly temperature and precipitation data of F-GSWP2-B1. Köppen's climate classification was then integrated into four climatic groups: tropical rain forest (Af); tropical monsoon, savanna, and dry climates (Am, Aw, B); temperate and continental [warmer] climates (C, Da, Db); and continental [cooler] and polar climates (Dc, Dd, E). This grouping of Köppen's climate classification is similar to that proposed by Nijssen et al. (2001). Finally, a climatic

group was assigned to each validation basin according to the majority of grid cells within the basin. The plots corresponded well to Budyko's semi-empirical curve (solid line), and there was no clear relationship as to which climate classification a basin belonged (Fig. 3a). In some basins with a low aridity index, the plots exceeded the energy-constraint line. This results from the negative sensible heat flux in northern high latitudes and mountainous areas (Milly and Dunne, 2002).

We assumed that the precipitation and simulated net radiation of GSWP2-B1 were correct and only simulated runoff was biased. We re-calculated the evaporation to precipitation ratio ($E/P=(P-R/P)$ using the observed runoff of the GRDC and plotted it (Fig. 3b). The distribution of the plot was clearly different from Budyko's curve, but there were relationships between Köppen's climatic classification and Budyko's curve. In dry climates, the plots reached higher than Budyko's curve; on the contrary, in polar and cooler continental climates, the plots were generally lower. The plots of temperate and warmer continental climates were distributed over a broad range; dryer basins reached above Budyko's curve, whereas basins in tropical forests (e.g., Amazon) were near the curve.

To correct the bias in runoff, the parameters of subsurface flow τ and γ in Eq. (B12) were modified for the four climatic groups. The parameter τ is a time constant to set the daily maximum subsurface runoff. The parameter γ is a shape parameter to set the relationship between subsurface flow and soil moisture. The global default value that was used to draw Fig. 3a was 100 days for τ and 2.0 for γ. A series of simulations was conducted, shifting parameter τ from 50 to 300 at 50-day intervals and parameter γ from 0.5 to 3 at an interval of 0.5. Of 36 combinations, the combination of τ and γ, respectively, that minimized the sum of the root mean square errors of monthly stream-flow in each climatic group was determined: (100, 2.0) for tropical forest; (300, 2.0) for tropical monsoon, savanna, and dry climates; (200, 2.0) for temperate and continental [warmer] climates; and (50.0, 1.0) for continental [cooler] and polar climates. These modified parameter sets substantially improved the R-GSWP2-B1. In the following sections, only the results obtained using the modified parameters are shown for clarity

HESSD
4, 3535–3582, 2007

of discussion.

6.3 Routing of earlier studies

Three gridded global runoff data sets of earlier studies were also routed using the river routing module. R-D03, R-F02, and R-N01 were re-gridded so that they matched both the land/sea distribution and the spatial resolution of ISLSCP2. For N01 data, which have 2.0°×2.0° resolution, identical runoff was allocated to four 1.0°×1.0° grids. For R-F02 and R-D03 data, which have 0.5°×0.5° resolution, the runoff of four grids was aggregated into one grid. The Antarctic, Greenland, and lake grids (e.g., Great Lakes in USA and Canada) were excluded from analysis. Finally, simulated streamflow was obtained at 32 river gauging stations.

6.4 Sensitivity tests of the routing module

Because we fixed the flow velocity of the routing module at 0.5 m s^{-1} globally, a sensitivity test was conducted on the flow velocity of the routing module. Using R-GSWP2-B1 gridded runoff, 10-year (1986–1995) mean daily runoff was created. It was then routed using the river routing module and changing the flow velocity from 0.1 to 1.0 m s^{-1} at 0.1-m s^{-1} intervals. In this way, the arrival date of peak streamflow was obtained for each case. The best flow velocity was selected from 10 cases that produced an arrival date of peak streamflow close to the observed. Finally, the difference in the peak arrival date of streamflow between that best case and the 0.5-m s^{-1} case were listed for the five largest validation river basins in the world (Table 3). A faster flow velocity produces the faster arrival of a peak, and vice versa. The largest difference is observed at Stolb station in the Lena River (up to 34 days), but the remaining rivers differ by no more than 12 days. Considering that these basins are the largest basins in the world, so that the differences in the remaining basins are presumably smaller, we judged 0.5 m s^{-1} flow velocity to be acceptable. The hydrograph peak is primarily determined by the timing of runoff generation by the land surface hydrology module, and the flow velocity of the

HESSD
4, 3535–3582, 2007

routing module has limited effects on the results.

7 Results and discussion

7.1 Continental runoff

R-GSWP2-B1 was within the plausible range of runoff in Asia, North America, South America, Oceania, and the globe (Fig. 4). In Europe and Africa, it exceeds plausible values by 27% and 10%, respectively. R-GSWP1 were the smallest among the data sets in every continent, and simulated runoff was below the plausible range in all continents except Africa (Fig. 4). In contrast, R-GSWP2-B0 was the largest among the data sets. In this case, simulated runoff greatly exceeded the range of the three earlier projections in Europe and North America (Fig. 4).

We then identified the characteristics of the runoff simulation by the land surface hydrology module. Mean annual runoff simulated by the land surface hydrology module and the multi-model ensemble mean of GSWP participating LSMs were compared for both F-GSWP1 and F-GSWP2-B0. F-GSWP2-B1 was newly developed here and an ensemble mean is lacking; in addition, different LSMs participated in GSWP1 and GSWP2. Multimodel ensemble mean products have been found to perform significantly better than single-model systems in weather and seasonal climate forecasts, and Guo et al. (2007) applied the multimodel approach to the land surface component to improve the quality of simulated soil moisture. They found that the simple average of 17 soil moisture products of the GSWP2 participating model outperforms most individual products in simulating the phasing of the annual cycle, interannual variability, and the magnitude of observed soil moisture. The runoff simulated by the land surface hydrology module agrees well with that of the multimodel ensembles of state-of-the-art LSMs in mean annual runoff at the continental scale (Fig. 4). Discrepancies range from −15 to 10% for R-GSWP1 and from −9% to 6% for R-GSWP2-B0. Thus, the discrepancy among meteorological forcing inputs is larger than that among LSMs.

HESSD
4, 3535–3582, 2007

7.2 Relationship between precipitation and runoff

There was a clear linear relationship between precipitation and runoff (Fig. 5). The observation-based data sets, namely R-F02, R-D03, and R-BR75, gave similar results, indicating that they are consistent in precipitation and runoff. Except for a few cases, R-GSWP2-B0 projects into the upper right, which means larger input precipitation was used compared to the earlier studies, and larger runoff was produced; R-GSWP1 projects into the lower left (Fig. 5). This linear relationship emphasizes precipitation as the dominant factor in the production of continental runoff. Except for Europe and Africa, R-GSWP2-B1 precipitation is similar to that of R-F02, R-D03, and R-BR75; consequently, the projected runoff is also similar. The hydrological simulation of the land surface hydrology module shows better performance with F-GSWP2-B1 compared to F-GSWP1 and F-GSWP2-B0, mainly because its precipitation input is plausible.

The large precipitation in Europe from F-GSWP2-B1 produced large runoff (Fig. 5). This large precipitation was caused by the rain gauge undercatch correction applied to F-GSWP2-B1 precipitation forcing inputs. Even larger precipitation (similar to F-GSWP2-B1) was given; R-N01 is similar to that of R-F02, R-D03, and R-BR75. This is primarily a result of the basin-level hydrological parameter tuning applied to their model. Nijssen et al. (2001) showed two simulation results: one with parameter tuning (shown in Fig. 5) and one without. The simulated runoff in Europe with parameter tuning was 280 mm yr^{-1}; without parameter tuning, it was 333 mm yr^{-1}, which is similar to R-GSWP2-B1 (357 mm yr^{-1}). In Africa, R-GSWP2-B1 is large for precipitation input when a linear relationship is assumed between precipitation and runoff. The occurrence of large runoff is attributable to factors other than precipitation. This issue will be discussed further in the next subsection.

7.3 Continental zonal mean runoff

To analyze the distribution within each continent, we examined the continental zonal mean runoff for every 5° (latitude; Fig. 6). R-GSWP2-B1 reproduced continental runoff

HESSD
4, 3535–3582, 2007

in Asia, North America, South America, and Oceania quite well (Fig. 5); the zonal mean distribution also agrees well with that of earlier studies (Fig. 6). However, at lower latitudes in Asia, R-GSWP2-B1 is close to the upper limit of the range of the three earlier studies, and in Oceania, it is close to the lower limit. The runoff distribution in Africa indicates that the runoff at lower latitudes exceeds that of the previous studies. There is no significant difference in zonal mean precipitation among earlier studies and F-GSWP2-B1 (not shown). In the land surface hydrology module, evaporation is basically correlated with wind speed (Appendix B). The F-GSWP2-B1 wind speed is low at low latitudes, and evaporation is considered to be restricted (Fig. 1). These two phenomena seem to be key factors. In Europe, R-GSWP2-B1 constantly overestimated runoff by approximately 100–200 mm yr^{-1} in each zone. There was no discrepancy in simulated runoff between the land hydrology module and the multimodel ensemble mean in Europe (Fig. 4); therefore, overestimation does not seem to be caused by the land surface hydrology module. It may be argued that the overestimation can be attributed to the omission of anthropogenic withdrawals. However, Shiklomanov (2000) reported the annual total consumptive withdrawal in Europe as 189 km^3 yr^{-1}, or 19 mm yr^{-1}, which is too small to explain the overestimation.

7.4 Runoff in individual basins

The simulated streamflow data sets were validated at 32 major river gauging stations. Using the simulated and observed data, the normalized bias of mean annual streamflow (NBIAS), the difference in the arrival of peak streamflow (PEAK), and the correlation coefficient (*CC*) of annual streamflow variation were calculated as follows:

$$\mathrm{NBIAS} = (\bar{s} - \bar{o}) / \bar{o} \tag{4}$$

$$\mathrm{PEAK} = \sum_{y=1986}^{1995} \left| m_{y,\mathrm{sim}} - m_{y,\mathrm{obs}} \right| \tag{5}$$

HESSD
4, 3535–3582, 2007

$$CC = \frac{\sum_{y=1986}^{1995} (s_y - \bar{s})(o_y - \bar{o})}{\sqrt{\sum_{y=1986}^{1995} (s_y - \bar{s})^2} \sqrt{\sum_{y=1986}^{1995} (o_y - \bar{o})^2}} \quad (6)$$

where $\bar{o}$ is the mean annual observed streamflow (calculated from available records between 1986 and 1995), $\bar{s}$ is the mean annual simulated streamflow (calculated for months in which $\bar{o}$ was available), $m_{y,\mathrm{sim}}$ is the month in which the simulated monthly hydrograph recorded the maximum streamflow, $m_{y,\mathrm{obs}}$ is the month of observation, s_y is the monthly simulated streamflow, and o_y is the monthly observed streamflow. The subscript y indicates the year. NBIAS is calculated to evaluate the simulated water balance in basins, PEAK to evaluate the timing of streamflow peaks in basins, and *CC* to evaluate the interannual variation in streamflow. Because lakes and reservoirs affect PEAK and *CC* considerably, six river basins, namely, Don, Parana, Sao Fransisco, Colorado, St. Lawrence, and Nelson, were excluded from calculations of PEAK and *CC*. As far as we know, the global gridded runoff data of R-F02 (Fekete et al., 2002) and that of R-D03 (Döll et al., 2003) are generally regarded as the best available data. Our focus here was to examine how closely our results fit these previous data. R-BR75 was excluded because it reports only zonal mean runoff data.

First, we examine NBIAS (Fig. 7a). For R-GSWP1, runoff in most basins in the northern middle to high latitudes is underestimated, except in Europe. In contrast, for R-GSWP2-B0, the runoff in a large number of basins in North America, Europe, and western Siberia was significantly overestimated (>50% of observed); NBIAS in eastern Siberia (e.g., Amur River and Lena River) was an exception, as the simulated runoff was well reproduced. These results are similar to the findings of Decharme and Douville (2006). For R-GSWP2-B1, NBIAS is less than ±50%, except for some river basins in Africa and northeastern South America. These basins are located in semi-arid to arid climatic zones (see Fig. 3b), and in these basins, runoff was significantly overestimated (>50% and sometimes >100% of the mean annual difference). The runoff of these basins was commonly overestimated in most of the earlier studies. R-N01 reproduced runoff fairly well, especially for river basins in Siberia. R-F02 and

R-D03 showed even better agreement, although, these two results are not surprising because the data sets were scaled so that simulated streamflow matched long-term mean annual streamflow. There were some basins with errors >20% because the period selected for scaling in these studies may have differed from ours.

Second, we examine PEAK (Fig. 7b). R-GSWP2-B1 reproduced the timing of long-term mean monthly streamflow well. In most of the basins, the error was within $\pm 2\,\mathrm{mo\,yr^{-1}}$. The results of the earlier studies R-N01, R-F02, and R-D03 are not shown because these data are at a monthly interval and thus are too coarse for a discussion of monthly peak flow.

Third, we examine *CC* (Fig. 7c). Because the simulation period of R-GSWP1 was only 2 years and that of R-F02 involved 1 year of climatological information, the *CC*s of these studies are not shown. As in Fig. 5, simulated runoff (or streamflow) was strongly correlated with input precipitation. Because precipitation in the earlier studies was based on ground observations, it seems that the annual variation in simulated runoff agrees well with the observations.

The land surface hydrology module is a submodel of an integrated global water resources model, which is intended for the assessment of the subannual temporal distribution of water resources and water use. Here, we set arbitrary thresholds for NBIAS, PEAK, and *CC*. The first threshold set was ±20% for NBIAS, ±1 mo for PEAK, and 0.8 for *CC*. The threshold for NBIAS was derived from Fig. 7a; the R-F02 and R-D03 data supported these criteria for most of the basins. The number of basins meeting these criteria was counted for each study (Table 4). The number of basins below the threshold of NBIAS clearly differed among the data sets. R-GSWP2-B1 showed far better performance than R-GSWP2-B0 and R-GSWP1 and was similar to R-N01; however, only 14 of the 32 river basins met the criteria. R-D03 was generated so that long-term mean annual discharges match, but not all river basins agreed with observations within ±20% error. The number of basins meeting the criteria of PEAK and *CC* were 16 and 13 of the 32 river basins, respectively. The performance of *CC* was quite similar among R-GSWP2-B1, R-N01, and R-D03. Our results indicate that it is still a challenge for

global hydrology models to simulate annual river discharge year by year. We changed the threshold set to ±50% for NBIAS, ±2 mo for PEAK, and 0.6 for *CC* (Table 4). In this case, NBIAS, PEAK, and *CC* of R-GSWP2-B1 agreed with the criteria for 24, 25, and 22 of 32 river basins, respectively. Because 70–80% of the validation basins fell within the criteria, we can state that these criteria indicate the simulation performance of the R-GSWP2-B1. In other words, water resource assessments should take the limited ability of the land surface hydrological module and the river routing model into account.

The goal of the integrated model was to assess the subannual distribution of water resources and water use. Figure 8 shows the normalized monthly streamflow of R-GSWP2-B1 and observations at 32 validation basins from 1986 to 1995. The monthly streamflow was normalized so that the mean annual streamflow from 1986 to 1995 equaled one. It is clear that significant seasonality occurs in many basins; these exhibited more than three times the mean annual streamflow for only a few months per year, and in the remaining months, streamflow was far less than one. The results indicate that we can move forward to assess the seasonal variability in global water resources.

8 Conclusions

To assess global water resources, an integrated model was developed consisting of six modules. In this report, meteorological forcing inputs and two natural hydrological cycle modules were presented and validated. GSWP1 and GSWP2 succeeded in producing comprehensive data sets of land surface flux and state variables involving dozens of state-of-the-art land surface models. However, Oki et al. (1999) and Decharme and Douville (2006) reported the shortcomings of runoff products from both of these projects. The products have strong biases compared to earlier studies (shown in Figs. 4, 6) and they poorly reproduce observations in individual basins (Fig. 7). One of the clear causes of the bias was the precipitation forcing input. The precipitation of F-GSWP1 was apparently lower than in earlier studies, whereas that of F-GSWP2-B0 was greater. Therefore, we revisited the original GSWP2 meteorological forcing

input (F-GSWP2-B0) and developed an improved meteorological forcing data set (F-GSWP2-B1). F-GSWP2-B1 had similar precipitation forcing inputs compared to reliable earlier studies, except in Europe. Using the F-GSWP2-B1 data and a newly developed land surface hydrology module designed to simulate surface energy and global runoff, global runoff was simulated for 10 years (1986–1995) at a resolution of $1^{\circ} \times 1^{\circ}$. In this simulation, both energy and water balance were closed in each grid cell. The simulation agreed well with earlier studies in the global, continental, and continental zonal mean distributions of runoff and reproduced 32 major river gauge stations in the world quite well, except in the semi-arid to arid climatic zones.

The input meteorological data and the land surface hydrology module are the basis of our global integrated water resources model. One of the aims of the integrated model is to assess the subannual temporal variation in water resources and water use. Overall the land surface hydrology model successfully reproduced the timing of peak streamflow and the shape of monthly hydrographs at river basins (Fig. 8). The application of the model to water resource assessments that take seasonal variability into account may be possible. However, because R-GSWP2-B1 showed substantial error in some basins, this requires caution, and extensive sensitivity tests must be conducted for the simulated runoff estimates.

The reproducibility of R-GSWP2-B1 is attributed to the meteorological forcing input of F-GSWP2-B1, but also relies on the parameter modifications discussed in Sect. 6.2. This parameter modification was based on findings that the bias in runoff is related to the climatic classifications proposed by Budyko and Köppen. These findings agree with those of Nijssen et al. (2001). Other than climate, soil and vegetation type might be tested to classify the hydrological parameters. Milly and Shmakin (2002) developed the LaD land surface model, which is based on the original bucket model (Manabe, 1969). In the LaD model, the parameters were set by soil and vegetation type, not by climatic classification.

During the validation process, a common distribution in the bias of mean annual streamflow was observed among this and earlier studies. A trend for underestimation

in northern high latitudes, especially in eastern Siberia, is one common feature. This problem was clearly observed in R-GSWP1 and R-GSWP2-B1. However, R-N01 and R-D03 also had this problem before tuning and correction. Döll et al. (2003) showed that unless a multiplying correction factor was applied, the R-D03 runoff in these regions was underestimated. Nijssen et al. (2001) showed that their land surface model produced quite small runoff in these areas using non-tuned (i.e., standard) parameter sets. To cope with this problem, a rain gauge undercatch correction methodology proposed by Motoya et al. (2002) was applied in the preparation of F-GSWP2-B1 input precipitation. Indeed, the R-GSWP2-B1 performed better at northern high latitudes than did the GSWP1 without correction, but it resulted in significant overestimation in Europe and moderate overestimation in some parts of western Siberia. Döll et al. (2003) also compared simulated results with and without rain gauge undercatch correction. They found that the rain gauge corrected precipitation improved annual runoff estimates in snow-dominated basins mainly in Siberia, although it resulted in significant overestimation of runoff in other regions, especially in central Europe; therefore, they ceased to use the correction. Underestimations in runoff may be attributed not only to underestimations in precipitation, but also to problems in representing hydrological processes. Nijssen et al. (2001) argued that surface storage might be important in the arctic climate because snowmelt water cannot infiltrate frozen soil. Further detailed studies are required to pinpoint the mechanisms that are currently lacking in runoff simulations in arid and arctic river basins; hopefully, these can be supported by field observations. In addition, overestimations in arid basins were commonly observed among the studies.

We introduced the results of global hydrological simulation using a single land surface model and a simple river routing model. Estimations of global water balance have been reported in a number of earlier publications in recent decades. However, taking the increasing public awareness of the effects of climate change and sustainability of world water resources into account, the reproducibility and limitations of simulation results should be disclosed in global water resource assessments for purposes of validity and transparency.

This report contains information useful to GSWP2 participants and the land surface modeling community. Our results suggest that F-GSWP2-B1 significantly improves upon hydrological simulations compared to F-GSWP2-B0; therefore, GSWP2 participants are encouraged to re-run their model using it. This report validates the GSWP2 meteorological forcing inputs through the output product of runoff. Because forcing inputs significantly affect the performance of simulations and consequently the analyses of the GSWP2, the attempt to validate and improve the meteorological forcing input is an essential task of GSWP participants. Compared to the GSWP1 (first phase of the GSWP), the GSWP2 extends the simulation period from 2 years to 10 years and increases the temporal resolution from ten days to one day so that both meteorological forcing input and output land surface products of the GSWP2 have a wider range of potential applications. The GSWP2-B1 meteorological forcing input is available at (ftp://hydro.iis.u-tokyo.ac.jp/pub/hanasaki/gswp2). The GSWP2-B0 output products are available at the GSWP2 Data Inter-Comparison Center (http://haneda.tkl.iis.u-tokyo.ac.jp/gswp2), but the hydrological products should be used with caution as noted in our assessment.

Appendix A

F-GSWP2-B0 meteorological forcing input

The F-GSWP2-B0 meteorological forcing input is a hybrid product of NCEP-DOE reanalysis (Kanamitsu et al., 2002) and various observation-based monthly gridded meteorological data. The air temperature input is a hybrid product of NCEP-DOE reanalysis (Kanamitsu et al., 2002) and observation-based monthly temperature data of the Climate Research Unit at University of East Anglia (CRU; New et al., 2000). First, the air temperature from the original NCEP-DOE reanalysis was re-gridded from the native 1.9°×1.9° resolution to the ISLSCP2 required 1°×1° resolution and processed from hourly data to three-hourly data (Zhao and Dirmeyer, 2003). The air temperature

HESSD
4, 3535–3582, 2007

of the CRU was scaled to adjust for the altitude difference between the CRU grid and the ISLSCP2 mean altitude. The NCEP-DOE reanalyses were linearly scaled so that the monthly mean values were identical to the CRU values. The air temperature data were linearly scaled again so that the diurnal temperature range for each month was identical to that of the CRU. In this way, the air temperature of NCEP-DOE was linearly scaled so that the monthly maximum, minimum, and mean air temperatures were identical to those of the CRU. The daily and three-hourly variations were not corrected; they were determined by the NCEP-DOE reanalysis.

For specific humidity and air pressure, the original NCEP-DOE data were corrected so that they were consistent with the altitude of ISLSCP2 and air temperature. For wind speed, NCEP-DOE data were used without any correction. For longwave and shortwave downward radiation, three-hourly Surface Radiation Budget (SRB) data produced at the NASA/Langley Research Center were used (Stackhouse et al., 2000).

The precipitation forcing input is also a hybrid product of the NCEP-DOE reanalysis, the global observational data set GPCC (Rudolf et al., 1994), and the GPCP (Huffman et al., 1997). GPCC data were used for grids where rain gauges were densely located, and GPCP data were used for grids where they were sparsely located. For precipitation data, an algorithm for rain gauge correction for wind-caused undercatch was applied to the rainfall and snowfall input data set (Motoya et al., 2002). The method was described in Sect. 3.

Appendix B

The land surface hydrology module

B1 Albedo

The albedo scheme was identical to that of Robock et al. (1995). The snow-free albedo (α_{soil}) was taken from the GSWP2 standard monthly land use data set and included

plant phenological aspects. The snow surface albedo (α_{snow}) varied according to the surface temperature (T_s) as follows (Robock et al., 1995):

$$\alpha_{snow} = \begin{cases} \alpha_{max} & T_S \le T_{crit} \\ \{\alpha_{max} \times (T_{melt} - T_s) + \alpha_{min} \times (T_s - T_{crit})\} / (T_{melt} - T_{crit}) & T_{crit} \le T_S \le T_{melt} \\ \alpha_{min} & T_{melt} \le T_S \end{cases} \quad \text{(B1)}$$

where α_{max} is the maximum snow albedo, fixed at 0.60; α_{min} is the minimum snow albedo, fixed at 0.45; T_{crit} is the critical temperature (263.15 K); and T_{melt} is the melting point of ice (273.15 K). The surface albedo was expressed as:

$$\alpha = \begin{cases} \alpha_{snow} & 20 < \text{SWE} \\ \alpha_{soil} + \sqrt{0.05 \cdot \text{SWE}} \times (\alpha_{snow} - \alpha_{soil}) & 0 < \text{SWE} \le 20 \\ \alpha_{soil} & \text{SWE} = 0 \end{cases} \quad \text{(B2)}$$

where SWE is the snow water equivalent [kg m^{-2}].

B2 Sensible heat and latent heat

Potential evaporation E_P [kg m^{-2} s^{-1}] was expressed as:

$$E_P(T_S) = \rho C_D U (q_{SAT}(T_S) - q_a) \quad \text{(B3)}$$

where ρ is the density of air [kg m^{-3}], C_D is the bulk transfer coefficient (0.003), U is the wind speed [m s^{-1}], $q_{SAT}(T_S)$ is the saturated specific humidity at surface temperature [kg kg^{-1}], and q_S is the specific humidity [kg kg^{-1}]. Evaporation from a surface (E) was expressed as:

$$E = \beta E_P(T_S) \quad \text{(B4)}$$

where

$$\beta = \begin{cases} 1 & 0.75 W_f \le W \\ W / W_C & W < 0.75 W_f \end{cases} \quad \text{(B5)}$$

HESSD
4, 3535–3582, 2007

where W is the soil water content [kg m^{-2}] and W_f is the soil water content at field capacity, which was fixed at 150 [kg m^{-2}]. The sensible heat flux (H) is expressed as:

$$H = C_p^* \rho C_D U (T_s - T_a) \tag{B6}$$

where C_p^* is the specific heat of air [1005 J kg^{-1} K^{-1}] and T_a is the air temperature [K].

B3 Energy balance

The energy balance was expressed as:

$$(1 - \alpha) SW^{\downarrow} + LW^{\downarrow} = \sigma T_S^4 + \iota E + H + G \tag{B7}$$

where $SW^{\downarrow}$ is the downward shortwave radiation, $LW^{\downarrow}$ is the downward longwave radiation, σ is the Stefan-Boltzmann constant, ι is the latent heat [2.45×10^6 J kg^{-1}], and G is the soil heat flux. The original bucket model (Manabe, 1969) does not have soil heat capacity or soil heat flux because it was not designed to simulate diurnal cycles; however, the meteorological forcing input of GSWP2 is three-hourly. We added the force restore method (Bhumralkar, 1975; Deardorff, 1978) to simulate the surface temperature:

$$\begin{cases} C_s \frac{\partial T_s}{\partial t} = (1 - \alpha) S^{\downarrow} + L^{\downarrow} - \sigma T_s^4 - \iota E - H - \omega C_S (T_s - T_d) \\ C_d \frac{\partial T_d}{\partial t} = (1 - \alpha) S^{\downarrow} + L^{\downarrow} - \sigma T_s^4 - \iota E - H \end{cases} \tag{B8}$$

where C_s is the surface heat capacity, C_d is the deep soil heat capacity ($C_d = \sqrt{365} C_S$ [J m^{-2} K^{-1}]), and ω is the angular velocity ($\omega = 2\pi / 24 \cdot 60 \cdot 60$ [s^{-1}]).

B4 Snow and soil water balance

The snow balance was expressed as:

$$\frac{\mathrm{dSWE}}{dt} = \mathrm{Snowf} - Q_{sm} - Q_{\mathrm{sub}} \tag{B9}$$

where SWE is the snow water equivalent [kg m^{-2}], Snowf is the snowfall rate [kg m^{-2} s^{-1}], Q_{sm} is the snow melt rate [kg m^{-2} s^{-1}], and Q_{sb} is the sublimation rate [kg m^{-2} s^{-1}]. With a snow-covered surface, soil moisture does not change through precipitation or evaporation. The soil water balance was expressed as:

$$\frac{dW}{dt} = \text{Rainf} + Q_{sm} - E - Q_s - Q_{sb} \tag{B10}$$

where Q_s is the surface runoff and Q_{sb} is the subsurface runoff [kg m^{-2} s^{-1}].

B5 Runoff

Surface runoff (Q_s) was generated if the soil water content exceeded the capacity of soil water (i.e., field capacity):

$$Q_s = \begin{cases} W - W_f & W_f < W \\ 0 & W \leq W_f \end{cases} \tag{B11}$$

Subsurface runoff (Q_{sb}), which was not in the original bucket model (Manabe, 1969; Robock et al., 1995), was incorporated to the model as:

$$Q_{sb} = \frac{W_f}{\tau} \left(\frac{W}{W_f} \right)^{\gamma} \tag{B12}$$

where Q_{sb} is the subsurface runoff [kg m^{-2} s^{-1}] and τ is a time constant [s]. This equation is similar to the percolation rate of the LPJ model (Gerten et al., 2004). The γ was set at 2, and τ at 100 days=86 400×365×100 s, both are globally constant.

B6 Mosaic

The module has two separate soil moisture regimes within a grid: one for irrigated areas and one for non-irrigated areas. Identical meteorological forcing input is given to both of the land use types, but surface fluxes and state variables are calculated

independently. This scheme is not used in natural hydrological cycle simulations, but is used in natural-anthropogenic coupled simulations when irrigation is taken into account. The soil moisture in irrigated croplands is distinguished from that in other areas.

Acknowledgements. We thank the Global Runoff Data Centre for providing us with their valuable data. We are grateful to N.-D. Thanh, K. Takahashi, and Y. Masutomi for valuable comments. This study was supported by the Global Environmental Research Fund S-4 from the Ministry of the Environment, Japan, and a JSPS/Grant-in-Aid for Scientific Research (S)19106008.

References

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rosch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global model of water use and availability, Hydrol. Sci. J., 48, 317–337, 2003a.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rosch, T., and Siebert, S.: Global estimates of water withdrawals and availability under current and future "business-as-usual" conditions, Hydrol. Sci. J., 48, 339–348, 2003b.

Alcamo, J., Florke, M., and Marker, M.: Future long-term changes in global water resources driven by socio-economic and climatic changes, Hydrol. Sci. J., 52, 247-275, 2007.

Arnell, N. W.: Climate change and global water resources, Global Environmental Change-Human and Policy Dimensions, 9, S31–S49, 1999.

Arnell, N. W.: Climate change and global water resources: SRES emissions and socio-economic scenarios, Global Environmental Change-Human and Policy Dimensions, 14, 31–52, 2004.

Baumgartner, A. and Reichel, E.: The world water balance, Elsevier Scientific Publishing Company, 1975.

Betts, A. K. and Beljaars, A. C. M.: ECMWF ISLSCP-II near-surface dataset from ERA-40, ECMWF, Shinfield Park, Reading, ERA-40 Project Report Series 8, 2003.

Bhumralkar, C.: Numerical experiments on the computation of ground surface temperature in an atmospheric general circulation model, J. Appl. Meteorol., 14, 1246–1258, 1975.

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., and Gerten, D.:

Modelling the role of agriculture for the 20th century global terrestrial carbon balance, Global Change Biol., 13, 679–706, 2007.
Budyko, M. I.: Climate and Life, Elsevier, New York, 1974.
Döll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270, 105–134, 2003.
Deardorff, J. W.: Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, J. Geophys. Res., 83-C4, 1889–1903, 1978.
Decharme, B. and Douville, H.: Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations, Clim. Dynam., 27, 695–713, 2006.
Dirmeyer, P. A., Dolman, A. J., and Sato, N.: The pilot phase of the Global Soil Wetness Project, B. Am. Meteorol. Soc., 80, 851–878, 1999.
Dirmeyer, P. A., Gao, X. A., Zhao, M., Guo, Z. C., Oki, T. K., and Hanasaki, N.: GSWP-2 – Multimodel anlysis and implications for our perception of the land surface, B. Am. Meteorol. Soc., 87, 1381–1397, 2006.
Fekete, B. M., Vorosmarty, C. J., and Grabs, W.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, Global Biogeochem. Cycles, 16, 1042, doi:10.1029/1999GB001254, 2002.
Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., and Sitch, S.: Terrestrial vegetation and water balance – hydrological evaluation of a dynamic global vegetation model, J. Hydrol., 286, 249–270, 2004.
Gervois, S., de Noblet-Ducoudre, N., Viovy, N., Ciais, P., Brisson, N., Seguin, B., and Perrier, A.: Including croplands in a global biosphere model: methodology and evaluation at specific sites, Earth Interactions, 8–16, 2004.
Guo, Z. C. and Dirmeyer, P. A.: Evaluation of the Second Global Soil Wetness Project soil moisture simulations: 1. Intermodel comparison, J. Geophys. Res.-Atmos., 111, D22s02, doi:10.1029/2006JD007233, 2006.
Guo, Z. C., Dirmeyer, P. A., Hu, Z. Z., Gao, X., and Zhao, M.: Evaluation of the Second Global Soil Wetness Project soil moisture simulations: 2. Sensitivity to external meteorological forcing, J. Geophys. Res.-Atmos., 111, D22s03, doi:10.1029/2006JD007845, 2006.
Guo, Z. C., Dirmeyer, P. A., Gao, X., and Zhao, M.: Improving the quality of simulated soil moisture with a multi-model ensemble approach, Q. J. Roy. Meteor. Soc., 133, 731–747, 2007.
Haddeland, I., Skaugen, T., and Lettenmaier, D. P.: Anthropogenic impacts on continental

surface water fluxes, Geophys. Res. Lett., 33, L08406, doi:10.1029/2006GL026047, 2006.
Hall, F. G., de Colstoun, E. B., Collatz, G. J., Landis, D., Dirmeyer, P., Betts, A., Huffman, G. J., Bounoua, L., and Meeson, B.: ISLSCP Initiative II global data sets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies, J. Geophys. Res.-Atmos., 111, D22s01, doi:10.1029/2006JD007366, 2006.
Hanasaki, N., Kanae, S., Oki, T., and Shirakawa, N.: An integrated model for the assessment of global water resources – Part 2: Anthropogenic activities modules and assessments, Hydrol. Earth Syst. Sci. Discuss., 4, 3583–3626, 2007, http://www.hydrol-earth-syst-sci-discuss.net/4/3583/2007/.
Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schneider, U.: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset, B. Am. Meteor. Soc., 78, 5–20, 1997.
Jachner, S., Gerten, D., Rohwer, J., and Bondeau, A.: How much water is used in global irrigated and rainfed agriculture?, Geophys. Res. Abstr., 9, 03325, 2007.
Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP-DOE AMIP-II reanalysis (R-2), B. Am. Meteor. Soc., 83, 1631–1643, 2002.
Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, J. Geophys. Res.-Atmos., 99, 14 415–14 428, 1994.
Manabe, S.: Climate and the ocean circulation 1. The atmospheric circulation and the hydrology of the earth's surface, Mon. Weather Rev., 97-11, 739–774, 1969.
McKnight, T. L. and Hess, D.: Climate zones and types, in: Physical geography: a landscape appreciation, edited by: McKnight, T. L. and Hess, D., Prentice Hall, Upper Saddle River, NJ, 200–240, 2000.
Meeson, B. W., Corprew, F. E., McManus, J. M. P., Myers, D. M., Closs, J. W., Sun, K.-J., Sunday, D. J., and Sellers, P. J.: ISLSCP Initiative I – Global data sets for land-atmosphere models, 1987–1988, NASA, 1995.
Milly, P. C. D. and Dunne, K. A.: Macroscale water fluxes – 2. Water and energy supply control of their interannual variability, Water Resour. Res., 38, 1206, doi:10.1029/2001WR000760, 2002.
Milly, P. C. D. and Shmakin, A. B.: Global modeling of land water and energy balances. Part I: the land dynamics (LaD) model, J. Hydromet., 3, 283–299, 2002.
Mo, X., Liu, S., Lin, Z., Xu, Y., Xiang, Y., and McVicar, T. R.: Prediction of crop yield, water

consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain, Ecol. Model., 183, 301–322, 2005.
Motoya, K., Masuda, K., Takata, K., and Oki, T.: Sensitivity of precipitation gauge correction for the estimation of the global water balance, Eos, Trans. Amer. Geophys. Union, 83, AbstractH51A-0775, 2002.
New, M., Hulme, M., and Jones, P.: Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology, J. Climate, 12, 829–856, 1999.
New, M., Hulme, M., and Jones, P.: Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate, J. Climate, 13, 2217–2238, 2000.
Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F.: Predicting the discharge of global rivers, J. Climate, 14, 3307–3323, 2001.
Oki, T. and Sud, Y. C.: Design of total runoff integrating pathways TRIP – A global river channel network, Earth Interactions, 2, EI013, doi:10.1175/1087-3562(1998)002<0001:DOTRIP>2.3.CO;2, 1998.
Oki, T., Nishimura, T., and Dirmeyer, P.: Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP), J. Meteorol. Soc. Jpn., 77, 235–255, 1999.
Oki, T., Agata, Y., Kanae, S., Saruhashi, T., Yang, D. W., and Musiake, K.: Global assessment of current water resources using total runoff integrating pathways, Hydrol. Sci. J., 46, 983–995, 2001.
Oki, T., Agata, Y., Kanae, S., Saruhashi, T., and Musiake, K.: Global Water Resources Assessment under Climatic Change in 2050 using TRIP, IAHS Publication, 280, 124–133, 2003.
Oki, T. and Kanae, S.: Global hydrological cycles and world water resources, Science, 313, 1068–1072, 2006.
Osborne, T. M., Lawrence, D. M., Challinor, A. J., Slingo, J. M., and Wheeler, T. R.: Development and assessment of a coupled crop-climate model, Global Change Biology, 13, 169–183, 2007.
Robock, A., Vinnikov, K. Y., Schlosser, C. A., Speranskaya, N. A., and Xue, Y. K.: Use of mid-latitude soil-moisture and meteorological observations to validate soil-moisture simulations with biosphere and bucket Models, J. Climate, 8, 15–35, 1995.
Rudolf, B., Hauschild, H., Rueth, W., and Schneider, U.: Terrestrial Precipitation Analysis: Operational Method and Required Density of Point Measurements, in: Global Precipitations

and Climate Change, edited by: Desbois, M. and Desalmand, F., NATO ASI I/26, Springer Verlag, Berlin, 173–186, 1994.
Sellers, P. J., Meeson, B. W., Closs, J., Collatz, J., Corprew, F., Dazlich, D., Hall, F. G., Kerr, Y., Koster, R., Los, S., Mitchell, K., McManus, J., Myers, D., Sun, K. J., and Try, P.: The ISLSCP initiative I global datasets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies, B. Am. Meteor. Soc., 77, 1987–2005, 1996.
Sevruk, B.: Method of correction for systematic error in point precipitation measurement for operational use, World Meteorological Organization, Geneva, pp. 91, 1982.
Sevruk, B. and Hamon, W. R.: International comparison of national precipitation gauges with a reference pit gauge, World Meteorological Organization, Geneva, WMO/CIMO Instruments and Observing Methods Rep. 17, 135, 1984.
Shiklomanov, I. A.: Appraisal and assessment of world water resources, Water Int., 25, 11–32, 2000.
Stackhouse Jr., P. W., Gupta, S. K., Cox, S. J., Chiacchio, M., and Mikovitz, J. C.: The SRB Project release 2 data set: An update., GEWEX News, 10-3, 2000.
Tanaka, K., Yorozu, K., Hamabe, R., and Ikebuchi, S.: Validation of the GSWP2 baseline simulation, American Meteorological Society 19th Conf. Hydrology, Paper 6.2, 2005.
Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B.: Global water resources: Vulnerability from climate change acid population growth, Science, 289, 284–288, 2000.
Willmot and collaborator's global climate resource pages: http://climate.geog.udel.edu/~climate/index.shtml, access: 15 September 2007.
Yamazaki, T.: A one-dimensional land surface model adaptable to intensely cold regions and its application in Eastern Siberia, J. Meteorol. Soc. Jpn., 79, 1107–1118, 2001.
Zhao, M. and Dirmeyer, P. A.: Production and analysis of GSWP-2 near-surface meteorology data sets, Center for Ocean-Land-Atmosphere Studies, COLA Technical Report 159, 2003.

Table 1. Differences in meteorological forcing inputs between F-GSWP2-B0 and F-GSWP2-B1.

Meteorological forcing	F-GSWP2-B0		F-GSWP2-B1	
	Diurnal/Synoptic	Annual/Monthly	Diurnal/Synoptic	Annual/Monthly
Precipitation	NCEP-DOE	GPCC and GPCP[a]	ERA40	GPCC[b]
Rain gauge undercatch correction algorithm	Motoya et al. (2002)[c]		Motoya et al. (2002)[d]	
Rain/snow separation algorithm	Original NCEP/DOE separation		Yamazaki (2001)	
Air temperature	NCEP-DOE and CRU[e]	CRU	ERA40	CRU
Specific humidity	NCEP-DOE[f]		ERA40[f]	
Air pressure	NCEP-DOE[g]		ERA40[g]	
Wind speed	NCEP-DOE		ERA40	
Shortwave radiation	NASA Langley Research Center Surface Radiation Budget Ver 2			
Longwave radiation				

[a]GPCC was used for grids where rain gauges were densely located, whereas GPCP was used for grids where they were sparsely located. [b]GPCP was not used. [c]The algorithm of Motoya et al. (2002) and NCEP-DOE wind speed at the height of 10 m was used. [d]The algorithm of Motoya et al. (2002) and ERA40 windspeed at the height of 2 m (originally 10 m) was used. [e]Daily maximum and minimum temperature changes were scaled linearly by CRU data. [f]Adjusted to corrected air temperature so that the relative humidity of the original reanalysis was conserved. [g]Adjusted to ISLSCP2 elevation.

HESSD
4, 3535–3582, 2007

Table 2. Earlier studies of global runoff estimation.

Data	Period	Time	Space	Source	Output Corr.[1]	Param. Tune[2]	Simulation time step	Precipitation
R-GSWP2-B0	1986–1995	Daily	1.0°×1.0°	Dirmeyer et al. (2006)			3 h	Rudolf et al. (1994) and Huffman et al. (1997)
R-GSWP2-B1	1986–1995	Daily	1.0°×1.0°	This study			3 h	Rudolf et al. (1994)
R-GSWP1	1987–1988	Daily	1.0°×1.0°	Oki et al. (1999)			6 h	Meeson et al. (1995)
R-N01	1980–1993	Daily	1.0°×1.0°	Nijssen et al. (2001)		Y	Day	Huffman et al. (1997)
R-F02	Clim.	Monthly	0.5°×0.5°	Fekete et al. (2002)	Y		Month	Willmott et al. (2007)
R-D03	1901–1995	Monthly	0.5°×0.5°	Döll et al. (2003)	Y	Y	Day	New et al. (2000)
R-BR75	Clim.	Annually	5.0° zonal	Baumgartner andReichel (1975)			Year	Original

[1] Output runoff data were corrected so that simulated long-term mean annual streamflow agreed with the observations.
[2] Model parameter was tuned at specific river basins.

Table 3. Sensitivity of flow velocity to the arrival of peaks in five large basins around the world. Optimal flow velocity shows the flow velocity producing the peak closest to the observation.

Station	River	Optimal flow velocity [$m\,s^{-1}$]	Difference in arrival of peak from 0.5 $m\,s^{-1}$ flow velocity [d]
Obidos	Amazon	0.4	8
Vicksburg	Mississippi	0.9	11
Salekhard	Ob	0.5	0
Stolb	Lena	0.9	34
Igarka	Yenisei	0.7	12

HESSD
4, 3535–3582, 2007

Table 4. The number of river gauging stations meeting the criteria for the normalized bias of mean annual streamflow (NBIAS), the difference in the month of arrival of the peak of streamflow [mo yr^{-1}] (PEAK), and the correlation coefficient of interannual streamflow variation (*CC*).

Data	$-0.5 \leq$NBIAS≤ 0.5	PEAK≤ 2.0	$0.6 \leq CC$	$-0.2 \leq$NBIAS≤ 0.2	PEAK≤ 1.0	$0.8 \leq CC$
R-GSWP1	14	23	–	3	10	–
R-GSWP2-B0	16	25	21	7	19	10
R-GSWP2-B1	24	25	22	14	16	13
R-N01	22	–	22	14	–	15
R-F02	30	–	–	25	–	–
R-D03	27	–	22	19	–	13
Total validation basins	32	27	27	32	27	27

HESSD
4, 3535–3582, 2007

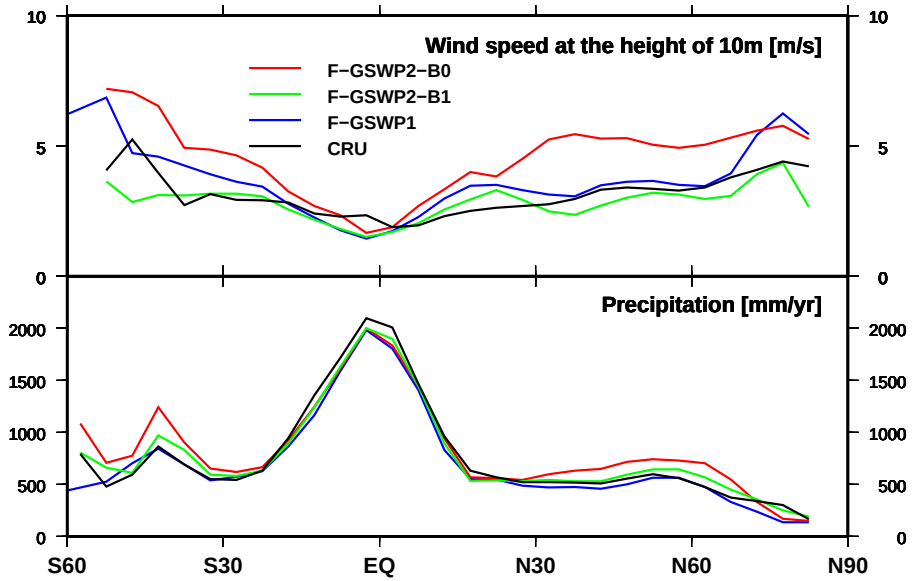

Fig. 1. Comparison of zonal mean wind speed and precipitation.

HESSD
4, 3535–3582, 2007

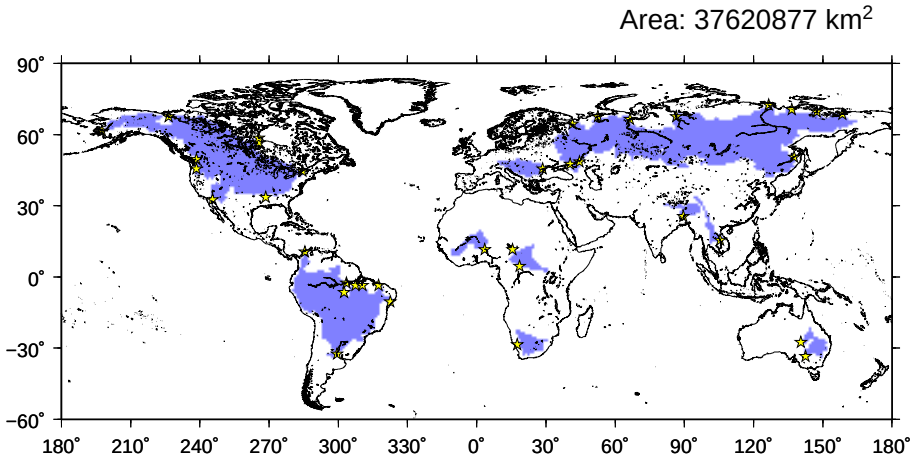

Fig. 2. Distribution of river gauging stations (stars). The shaded areas represent catchments.

HESSD
4, 3535–3582, 2007

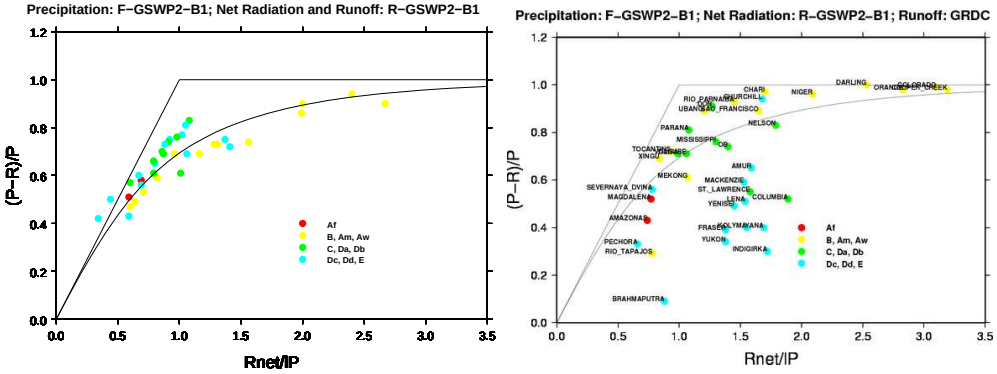

Fig. 3. Budyko's diagram for 37 basins (including five semi-arid and arid river basins that were excluded from the validation). **(a)** Simulated results of the land surface hydrology module forced using F-GSWP2-B1. R, mean annual runoff of the basin; P, precipitation; R_{net}, net radiation; solid line, energy and water constraint; curve, Budyko's semi-empirical curve of the relationship between $(P-R)/P$ and $R_{\mathrm{net}}/\iota P$. The color of plot indicates the Köppen climate classification: red, tropical rain forest (Af); yellow, tropical monsoon (Am) and savanna (Aw); green, temperate (C), hot summer continental (Da), and warm summer continental (Db); cyan, continental subarctic (Dc), continental subarctic with extreme severe winters (Dd), and polar (E). **(b)** Same as in (a), but runoff (R) was substituted by the GRDC observations.

HESSD
4, 3535–3582, 2007

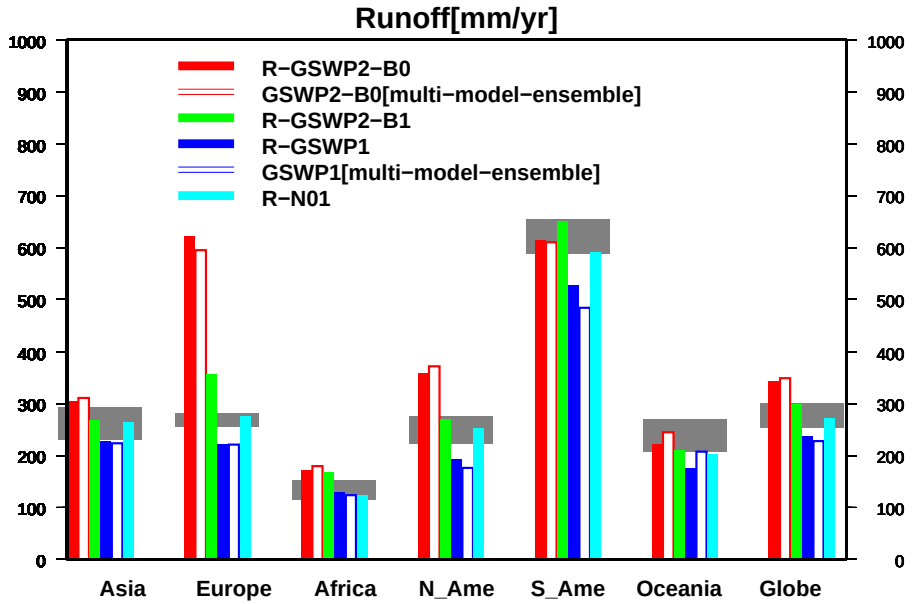

Fig. 4. Mean annual runoff for each continent. Gray shading indicates the range of runoff estimated by earlier observation-based studies (Baumgartner and Reichel, 1975; Fekete, 2002; Döll et al., 2003) and "multi-model-ensemble" indicates the results of the ensemble mean of state-of-the art land surface models participating in the GSWP. If the bar height lies within the shade, then runoff can be considered plausible.

HESSD
4, 3535–3582, 2007

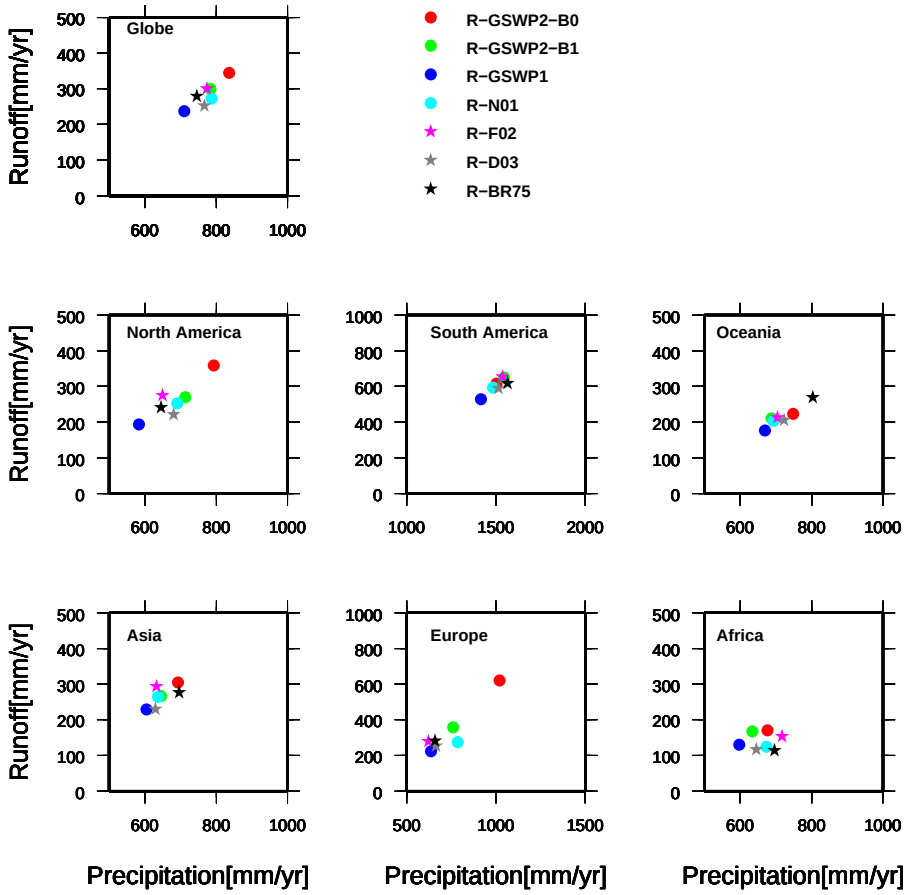

Fig. 5. The relationship between runoff and precipitation. Stars indicate observation-based studies.

HESSD
4, 3535–3582, 2007

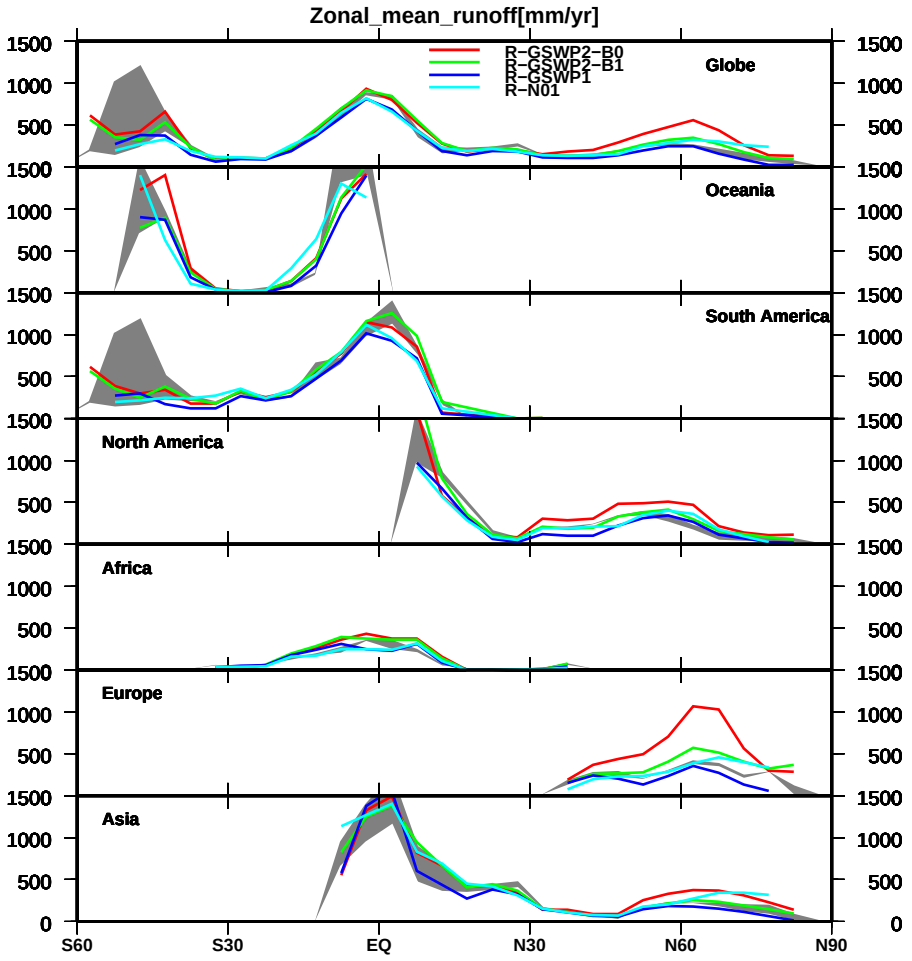

Fig. 6. Zonal mean runoff for six continents and the globe. Runoff was averaged for every five degrees of latitude. Gray shading indicates the range of runoff estimated by earlier observation-based studies as in Fig. 4.

HESSD
4, 3535–3582, 2007

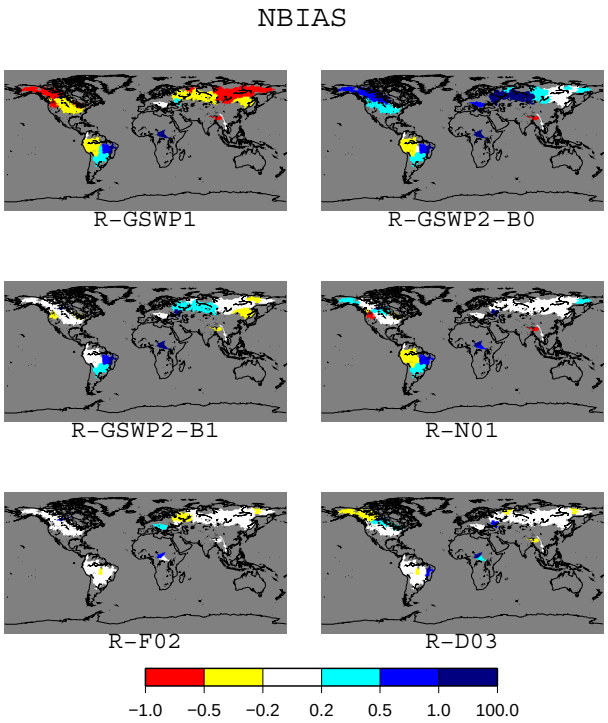

Fig. 7. Validation results for 32 river basins. **(a)** Normalized bias of mean annual runoff (NBIAS). **(b)** Delay in the arrival of peak streamflow (PEAK). R-N01, R-F02, and R-D03 are not shown because their data are monthly and thus too coarse for routing. **(c)** Correlation coefficient of annual streamflow (*CC*). R-GSWP1 and R-F02 are not shown because their simulation periods were 2 years and one climatological year, respectively.

HESSD
4, 3535–3582, 2007

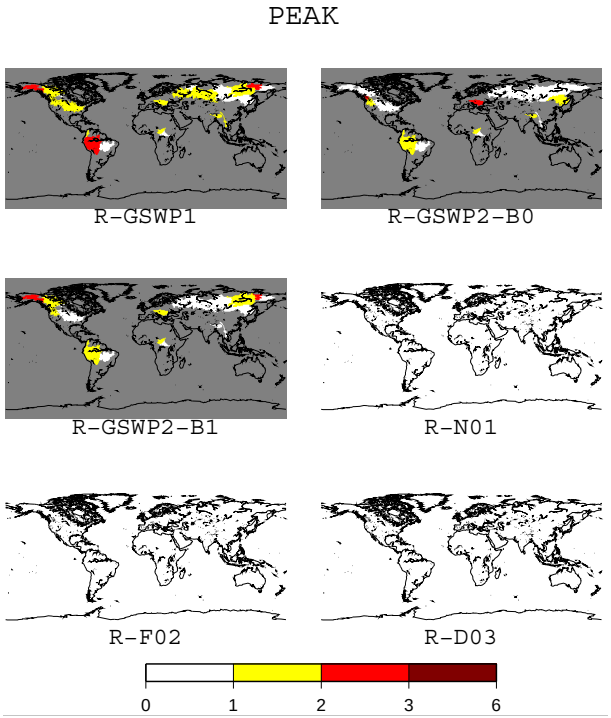

Fig. 7. Continued.

HESSD
4, 3535–3582, 2007

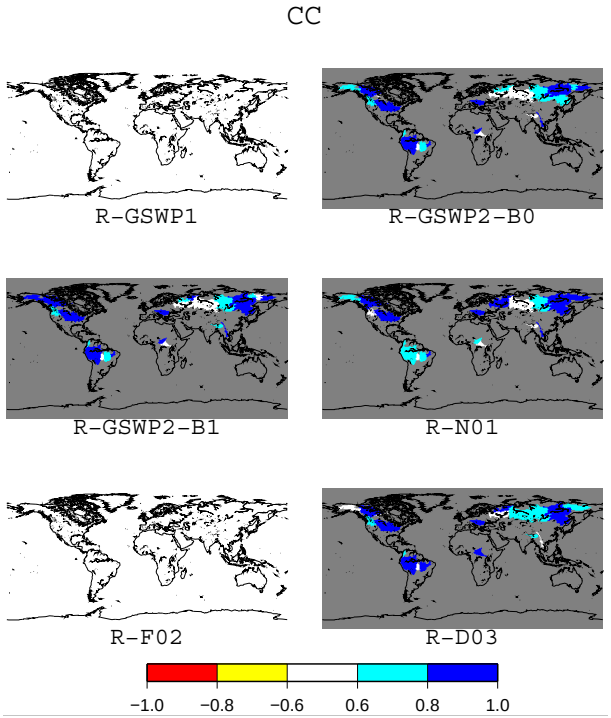

Fig. 7. Continued.

HESSD
4, 3535–3582, 2007

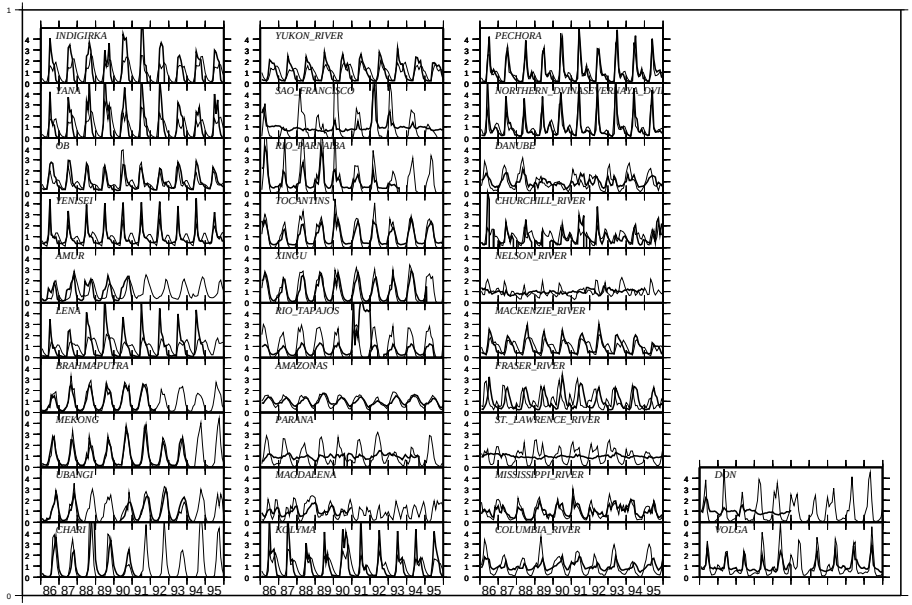

Fig. 8. Normalized monthly streamflow at 32 validation basins. Heavy line, observation; thin line, simulation. Monthly streamflow was normalized so that the mean annual streamflow from 1986 to 1995 (or available records in this period) equaled one.

HESSD
4, 3535–3582, 2007